

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 1, 2005, 13:15:28 ; Search time 35 Seconds
(without alignments)
3183.400 Million cell updates/sec

Title: US-10-659-782A-11
Perfect score: 1030
Sequence: 1 actctgagtggtgctgtt.....tggcagcagagaggtggtggtg 579

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlpl
-O=/cgn2 1/USPTO spoal_p/US10659782/runat_01022005 130352 14264/app query.fasta_1.775
-DB=PIR 79 -QFMT=fastan -SUFFIX=rrp -MINMATCH=0.1 -LOOPEXT=0
-UNITS=BITS -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODES=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10659782 @CEN 1 1 63 @runat 01022005 130352 14264 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	326	31.7	117	1 A59316	ghrelin precursor
2	256	24.9	117	1 B59316	ghrelin precursor
3	112	10.9	1344	1 A35175	mucin 1 precursor
4	105	9.8	152	2 T34649	hypothetical prote
5	103.5	10.0	379	2 T05441	proline-rich prote
6	103.5	10.0	379	2 D85257	extensin-like prot
7	100	9.7	620	2 S06733	hydroxyproline-ric
8	99.5	9.7	317	2 A28996	proline-rich prote
9	99.5	9.7	381	2 S16506	hypothetical prote
10	99	9.6	383	2 S32975	gene BCRF2 protein
11	98.5	9.6	349	2 T05857	hypothetical prote
12	98.5	9.2	356	1 WJHU2H	homeotic protein H
13	98	9.5	347	1 S10571	mucin 1 precursor,
14	96.5	9.4	506	2 B56201	transcription fact

15	96.5	9.4	514	2 A56201	transcription fact
16	96	9.3	839	2 F75518	hypothetical prote
17	95	9.2	632	2 T00084	hypothetical prote
18	94.5	9.2	1690	1 CGHUIB	collagen alpha 4(I
19	94	8.8	1216	2 T34101	hypothetical prote
20	94	9.1	3164	1 WMBEH6	UL36 protein - hum
21	93.5	9.1	349	2 JCS881	myocyte enhancer f
22	93.5	9.1	672	2 I40333	tracheal colonizat
23	93	9.0	1575	2 S68448	synaptojanin, 170K
24	92.5	8.7	454	2 T35380	probable membrane
25	92.5	8.7	591	2 B83039	probable thiol-dis
26	92	8.7	135	2 S12549	hypothetical prote
27	92	8.9	374	2 A42270	alpha (1,3) fucosy
28	92	8.9	544	2 S15664	transcription fact
29	91.5	8.9	511	1 A48560	UL54 protein - hum
30	91.5	8.9	512	1 WMBEY4	UL54 protein - hum
31	91.5	8.9	729	2 E70803	hypothetical prote
32	91.5	8.9	1006	2 T42731	atrophin-1 related
33	91.5	8.9	1091	2 S33596	protein-tyrosine k
34	91.5	8.6	1232	2 I38496	anion exchanger 3
35	91	8.8	240	2 A24264	proline-rich prote
36	91	8.8	539	2 T28770	hypothetical prote
37	91	8.8	576	2 T36729	probable serine/th
38	91	8.8	901	2 A49227	sialidase - Actino
39	90.5	8.8	339	2 JCS882	myocyte enhancer f
40	90.5	8.5	468	2 S26741	T-cell glycoprotei
41	90.5	8.8	587	2 T41653	probable transcrip
42	90	8.4	431	2 S09824	hypothetical prote
43	90	8.7	1184	2 G01763	atrophin-1 - human
44	90	8.7	1734	2 A54602	microtubule-associ
45	90	8.7	1794	2 T38459	hypothetical diver

ALIGNMENTS

RESULT 1
A59316
ghrelin precursor - human
N/Alternate names: preproghrelin
C/Species: Homo sapiens (man)
C/Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 09-Jul-2004
C/Accession: A59316
R/Kojima, M.; Hosoda, H.; Date, Y.; Nakazato, M.; Matsuo, H.; Kangawa, K.
Nature 402, 656-660, 1999
A/Title: Ghrelin is a growth-hormone-releasing acylated peptide from stomach.
A/Reference number: A59316; MUID:20067959; PMID:10604470
A/Accession: A59316
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-117 <KOU>
A/Cross-references: UNIPROT:Q9UBU3; GB:AB029434; NID:g6691571; PIDN:BAA89371.1; PID:g665
A/Experimental source: Tissue stomach endocrine cells
A/Note: submitted to GenBank, June 1999
C/Comment: Ghrelin secreted by the stomach stimulates the release of somatotropin (growth hormone) from the hypothalamus.
C/Superfamily: motilin
C/Keywords: hormone; lipoprotein; stomach
F/1-23/Domain: signal sequence #status predicted <SIG>
F/24-51/Product: ghrelin #status predicted <MAT>
F/52-117/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F/26/Binding site: octanoate (Ser) (covalent) #status experimental

Alignment Scores:
Pred. No.: 8.21e-21 Length: 117
Score: 326.00 Matches: 74
Percent Similarity: 53.19% Conservative: 1
Best Local Similarity: 52.48% Mismatches: 0
Query Match: 31.65% Indels: 66
DB: 1 Gaps: 1

US-10-659-782A-11 (1-579) x A59316 (1-117)

QY 112 ATGCCCTCCAGGACCGTCTGCAGCCTCTGCTCGGCATGCTCTGGCTGGACTTG 171
|||||

```

Db      1 MetProSerProGlyThrValCysSerLeuLeuLeuLeuLeuLeuMetLeuTrpLeuAspLeu 20
QY      172 GCATGGCAGGCTCCAGCTTCCTGAGCCTGAGCCCTGAACACACAGAGAGTCCAGGTGAGACCTCCC 231
      |||
Db      21 AlaMetAlaGlySerSerPheLeuSerProGluHisGlnArgValGln--Gln----- 37
QY      232 CACAAGCCCCACATGTTGTTCCAGCCCTGCCACTTAGCAACACAGACTCTGTGTGACCTGGAG 291
      |||
Db      37 ----- 37
QY      292 CAGCAGCGCCATCTCTGGGCTTCAGTCTTCTCCAGAGCACAAGGACTCTGGGTCTGTAC 351
      |||
Db      37 ----- 37
QY      352 CTCACCTGTTTCTGGAAGACATGGGGCTTAGAGTCTTAACACAGACTGTTTCCCCCTTCC 411
      |||
Db      37 ----- 37
QY      412 AGCAGAGAAGAGTCTGAAGAGCCACAGCAAGCTGCAGCCCGAGCTCTAGCAGGCT 471
      |||
Db      38 -----ArglysgluSerLysLysProAlaLysLeuGlnProArgAlaLeuAlaGlyT 56
QY      472 GCTCTCGCCCGGAGATGAGTCAAGCAGAGGGGAGAGGATGAAGTGAAGTCCGG 530
      |||
Db      56 rpleuargProGluaspGlyGlyGlnAlaGluGlyAlaGluaspGluLeuValArg 75

RESULT 2
B59316
ghrelin precursor - rat
N;Alternate names: preproghrelin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 09-Jul-2004
C;Accession: B59316
R;Kojima, M.; Hosoda, H.; Date, Y.; Nakazato, M.; Matsuo, H.; Kangawa, K.
Nature 402, 656-660, 1999
A;Title: Ghrelin is a growth-hormone-releasing acylated peptide from stomach.
A;Reference number: A59316; MUID:20067959; PMID:10604470
A;Accession: B59316
A;Status: not compared with conceptual translation
A;Molecule type: mRNA; protein
A;Residues: 1-117 <KO>
A;Cross-references: UNIPROT:Q9QVH7; GB:AB029433; NID:96691569; PIDN:BAA89370.1; PID:9669
A;Experimental source: strain SD; tissue stomach endocrine cells
A;Note: submitted to GenBank, June 1999
C;Comment: Ghrelin secreted by the stomach stimulates the release of somatotropin (growth hormone) from the hypothalamus.
C;Superfamily: motilin
C;Keywords: hormone; lipoprotein; stomach
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-51/Product: ghrelin #status predicted <MAT>
F;52-117/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;26/Binding site: octanoate (Ser) (covalent) #status experimental

Alignment Scores:
Pred. No.: 1.27e-14 Length: 117
Score: 256.00 Matches: 60
Percent Similarity: 46.81% Conservative: 6
Best Local Similarity: 42.55% Mismatches: 9
Query Match: 24.85% Indels: 66
Db: 1 Gaps: 1

US-10-659-782A-11 (1-579) x B59316 (1-117)
QY      112 ATGCCCTCCCGAGGACCGTCTGCAGCCTCTGCTCGGCATGCTCTGGCTGAGCTGTG 171
      |||
Db      1 MetValSerSerAlaThrIleCysSerLeuLeuLeuLeuSerMetLeuTrpMetAspMet 20
QY      172 GCCATGGCAGGCTCCAGCTTCCTGAGCCCTGAACACACAGAGAGTCCAGGTGAGACCTCCC 231
      |||
Db      21 AlaMetAlaGlySerSerPheLeuSerProGluHisGlnLys--Ala----- 35
QY      232 CACAAGCCCCACATGTTGTTCCAGCCCTGCCACTTAGCAACACAGACTCTGTGTGACCTGGAG 291
      |||
Db      35 ----- 35

```

```

QY      292 CAGCAGCGCCATCTCTGGGCTTCAGTCTTCTCCAGAGCACAAGGACTCTGGGTCTGTAC 351
      |||
Db      35 ----- 35
QY      352 CTCACCTGTTTCTGGAAGACATGGGGCTTAGAGTCTTAACACAGACTGTTTCCCCCTTCC 411
      |||
Db      36 -----G 36
QY      412 AGCAGAGAAGAGTCTGAAGAGCCACAGCAGCTGCAGCCCGAGCTCTAGCAGGCT 471
      |||
Db      36 lnglnArgLysGluSerLysLysProAlaLysLeuGlnProArgAlaLeuGluGlyT 56
QY      472 GCTCTCGCCCGGAGATGAGTCAAGCAGAGGGGAGAGGATGAAGTGAAGTCCGG 530
      |||
Db      56 rpleuHisProGluAspArgglyGlnAlaGluGlyAlaGluGluLeuGluLeuArg 75

RESULT 3
A35175
mucin 1 precursor, repetitive splice form A [validated] - human
N;Alternate names: breast carcinoma-associated DF3 antigen; core protein KP39; episialin
N;Contains: mucin 1 precursor, epithelial tumor antigen splice form; mucin 1 precursor,
C;Species: Homo sapiens (man)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Jun-2000
C;Accession: A35175; A35886; A35887; S10572; S40293; A36735; PX0066; S10218; S51
R;Ligtenberg, M.J.L.; Vog, H.L.; Gennissen, A.M.C.; Hilkens, J.
J. Biol. Chem. 265, 5573-5578, 1990
A;Title: Episialin, a carcinoma-associated mucin, is generated by a polymorphic gene enc
A;Reference number: A35175; MUID:90202794; PMID:2318825
A;Accession: A35175
A;Molecule type: mRNA
A;Residues: 1-952,1033-1344 <LIG>
A;Cross-references: GB:M32738; GB:J05288; NID:gl82121; PIDN:AAA35804.1; PID:gl82124; GB:
A;Experimental source: splice form A
A;Note: Genbank entries HUMEPISIA1 and HUMEPISIA2 present only the amino-and carboxyl-en
A;Accession: B35175
A;Molecule type: mRNA
A;Residues: 1-19,29-952,1033-1344 <LIG2>
A;Cross-references: GB:M32739; GB:J05288; NID:gl82126; PIDN:AAA35806.1; PID:gl82129; GB:
A;Experimental source: splice form B
A;Note: Genbank entries HUMEPISIA1 and HUMEPISIA2 present only the amino-and carboxyl-en
R;Gendler, S.J.; Lancaster, C.A.; Taylor-Papadimitriou, J.; Duhig, T.; Peat, N.; Burchel
J. Biol. Chem. 265, 15286-15293, 1990
A;Title: Molecular cloning and expression of human tumor-associated polymorphic epitheli
A;Reference number: A35886; MUID:90368715; PMID:1697589
A;Accession: A35886
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-19,29-952,1033-1344 <GEN>
A;Cross-references: GB:J05581; NID:gl88869; PIDN:AAA59876.1; PID:gl88870
A;Note: Genbank entry HUMMUCAB includes one copy of the tandemly repeated sequence
R;Lan, M.S.; Batra, S.K.; Qi, W.N.; Metzgar, R.S.; Hollingsworth, M.A.
J. Biol. Chem. 265, 15294-15299, 1990
A;Title: Cloning and sequencing of a human pancreatic tumor mucin cDNA.
A;Reference number: A35887; MUID:90368716; PMID:2394722
A;Accession: A35887
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-19,29-1109,'S',1111-1339,'A',1341-1344 <LAN>
A;Cross-references: GB:J05582; NID:gl89598; PIDN:AAA60019.1; PID:gl89599
R;Wreschner, D.H.; Hareuveni, M.; Tsarfaty, I.; Smorodinsky, N.; Horev, J.; Zaretsky, J.
Eur. J. Biochem. 189, 463-473, 1990
A;Title: Human epithelial tumor antigen cDNA sequences. Differential splicing may genera
A;Reference number: S10571; MUID:90276413; PMID:2351132
A;Accession: S10572
A;Molecule type: mRNA
A;Residues: 1-19,29-155,'P',157-175,'P',177-182,'A',184-212,1033-1344 <WRE>
A;Cross-references: EMBL:X52229; NID:G37053
R;Wreschner, D.H.
submitted to the EMBL Data Library, March 1990
A;Reference number: S40293

```

Db 153 -ThrArgProAlaProGlySerThrAlaProProAlaHisGlyValThrSerAlaProAs 172
 Qy 263 CACTTAGCAACCAAGCTCTGTGACCTGAGCAGCAGCGCCATC-----TCTGGCGCTTCAGTC 318
 Db 172 pThr--ArgProAla-----ProGlySerThrAlaProAlaHisGlyValThrSe 189
 Qy 319 TTCTCCAGAGCAAAAGGACTCTGGGTCTGACCTCACTCTGTTCTGGAAGACATGGGGG 378
 Db 189 rAlaProAspThrArgProAlaProGlySerThrAla-----ProProAlaHis--Gl 206
 Qy 379 CTTAGAGTCTCTAAACACACTGTTTCCCCCTTCCAGCAGAG-----AAAGG 423
 Db 206 yValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAlaHisGl 226
 Qy 424 AGTCGAAGAAGCCACCAGCAAGCTGAGCGCCCGAGCTCTAGCAGGCTGGCTCCGCCCGG 483
 Db 226 yValThrSerAlaProAspThrArgProAlaProGlySerThr-----AlaProAla 244
 Qy 484 AAGATGGAGTCAAGC-----AGAAGGGCGAGAGAGTGAAGTGGAAAGTCCCGG 531
 Db 244 aHisGlyValThrSerAlaProAspThrArgProAlaProGlySerThrAlaProAla 264
 Qy 532 TCGGTACTCTGCAGT 547
 Db 264 aHisGlyValThrSer 269
 RESULT 4
 T34649
 hypothetical protein SC1A11.02c SC1A11.02c - Streptomyces coelicolor (fragment)
 C/Species: Streptomyces coelicolor
 C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C/Accession: T34649
 R/Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandram
 submitted to the EMBL Data Library, January 1999
 A/Reference number: Z21551
 A/Accession: T34649
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-152 <SEE>
 A/Cross-references: EMBL:AL035205; PIDN:CAA22742.1; GSPDB:GN00070; SCODEB:SC1A11
 A/Experimental source: strain A3(2)
 C/Genetics:
 A/Gene: SCODEB:SC1A11.02c
 Alignment Scores:
 Pred. No.: 0.288 Length: 152
 Score: 105.00 Matches: 42
 Percent Similarity: 41.60% Conservative: 10
 Best Local Similarity: 33.60% Mismatches: 41
 Query Watch: 9.84% Indels: 32
 DB: 2 Gaps: 6
 US-10-659-782A-11 (1-579) x T34649 (1-152)
 Qy 372 TGTCCTTCAGAAACACTGAGGTGAGTCCAGACCAGAGTCTTTGTGCTCTGGGAGAGACTGA 313
 Db 1 CysProAla---SerSerProAlProProArg-----SerSerGly 13
 Qy 312 AGCCACAGATGGCGTCTGCTCCAGTCCAGTCCAGTGGTTGCTAAGTGGCAGG----- 258
 Db 14 ThrPro-ArgGlyAlaCysCysSerSerSerArgThrThrAlaArgGlyGlyArgCysPr 33
 -Qy 257 -----GCTGAAACAAATGTGGGGCTTTTGGGGAGGCTCTCAC 220
 Db 33 oAlaGlyArgSerSerProGlyThrGlyArgAlaArgGlyArgAlaArgGlyGlyArgPr 53
 Qy 219 CTGG-----ACTCTCTGGTGTTCAGGCTCAGGAAGCTCGAGC----- 182
 Db 53 oTrpArgArgSerAspSerThr-CysGlySerAlaGlyCysSerArgTrpThrGlyArgT 73
 Qy 181 -----CTGCCATGGCCCAAGTCCAGCAGATGCTCCGAGGAGGAGGCTGCACAGCGT 127
 Db 73 hrGlyArgAlaGlyProArgSerTrpArgThrCysThrThrAlaGlyCysCysArgArgT 93

QY 126 CCTGGGAGGCGCATGCCCTCAGCTGGGTTGACAGAGGTGGGCTGCAGTCTCGGCGGA 67
Db : : : : :
93 hrIleSerArg-----ArgSerGlyCysArg-ArgArgSerCysCysProGlyCys 109
QY 66 GTGGTGC 60
Db : : : : :
110 -TrpCys 111

RESULT 5

T05441
proline-rich protein F7K2.50 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05441
R;Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller
submitted to the Protein Sequence Database, November 1998
A;Reference number: Z15416
A;Accession: T05441
A;Molecule type: DNA
A;Residues: 1-379 <BEV>
A;Cross-references: UNIPROT:Q9SUX2; EMBL:AL033545
A;Experimental source: cultivar Columbia; BAC clone F7K2
C;Genetics:
A;Map position: 4
A;Note: F7K2.50
C;Superfamily: hydroxyproline-rich glycoprotein

Alignment Scores:
Pred. No.: 0.39 Length: 379
Score: 103.50 Matches: 55
Percent Similarity: 29.49% Conservative: 9
Best Local Similarity: 25.35% Mismatches: 71
Query Match: 10.05% Indels: 82
DB: 2 Gaps: 7

US-10-659-782A-11 (1-579) x T05441 (1-379)

QY 44 AGACCTGACAGACAGGACCCACTCCGACGAGTCCGAGGCTGCAGGCCACCTGTCTGCAACCCAG 103
Db : : : : :
52 GlnProAspProGlnProProThrProThrPheGlnProAlaProAlaAsnAsp 71
QY 104 CTGAGGCGCATGCCCTCCCGAGGACCGTCTGCAGCCCTCTGCTCCGCGATGCTCTGGC 163
Db : : : : :
72 GlnProProProProGlnSerThrSerProProValAlaThr----- 87
QY 164 TGGACTTGGCCATGGCAGGCTTCTGAGCCCTGAACACACAGAGAGTCCAGGTGA 223
Db : : : : :
88 ---ThrProAlaLeuProProLys-ProLeuPro---ProProLeuSerProPro 105
QY 224 GACCTCCCAAGCCCAACATGTTGTCAGCCCT----- 260
Db : : : : :
105 nThrThrProProProProAlaIleThrProProProProAlaIleThrProPr 125
QY 261 -----GCCACTTAGCAACAGC 277
Db : : : : :
125 oLeuSerProProProAlaIleThrProProProProAlaIleThrProProAl 145
QY 278 TCTGTGACCTGGAGCAGCGCCATCTCTGGGCTTCAGTCTTCTCCAGAGCACAAGG 337
Db : : : : :
145 aLeuProProLysProLeuProProLeu----- 155
QY 338 ACTCTGGGTCTGACCTCAGTCTTCTGGAAGGACATGGGGCTTAGAGTCTTAAACAGAC 397
Db : : : : :
156 -----SerProProGlnTh 160
QY 398 TGTTCCTCCCTTCCAGCAGAGAAGGAGTCGAAGAGCCACAGCCAGCTGCAGCCC-- 455
Db : : : : :
160 rThrProPro-----ProProProAlaIleThrProProLe 172
QY 456 -----CGAGCTCTAGCAGG 469
Db : : : : :
172 uSerProProLeuValGlyIleCysSerLysAsnAspThrGluLeuLysIleCysAlaG 192

QY 470 CTGGCTCCGCCCGAAGATGGA-----GGTCAAGCAGCAAGGGCAGAGGATGA 517
Db : : : : :
192 yIleLeuAlaIleSerAspGlyLeuLeuThrThrGlyArgAlaGluProCysCysSerIl 212
QY 518 ACTGGAAGTCCGGTCCGTACCTCTGCAGTCTTTATGCTTCTGTGGCAGC 566
Db : : : : :
212 eIleArgAsnValSerAspLeuAspAlaValThrCysPheCysLysSer 228

RESULT 6

D85257
extensin-like protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85257
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: D85257
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-379 <STO>
A;Cross-references: UNIPROT:Q9SUX2; GB:NC_001268; NID:g7369093; PIDN:CAB79202.1; GSPDB:C
C;Genetics:
A;Gene: AT4G22470
A;Map position: 4
C;Superfamily: hydroxyproline-rich glycoprotein

Alignment Scores:
Pred. No.: 0.39 Length: 379
Score: 103.50 Matches: 55
Percent Similarity: 29.49% Conservative: 9
Best Local Similarity: 25.35% Mismatches: 71
Query Match: 10.05% Indels: 82
DB: 2 Gaps: 7

US-10-659-782A-11 (1-579) x D85257 (1-379)

QY 44 AGACCTGACAGACAGGACCCACTCCGACGAGTCCGAGGCTGCAGGCCACCTGTCTGCAACCCAG 103
Db : : : : :
52 GlnProAspProGlnProProThrProThrPheGlnProAlaProAlaAsnAsp 71
QY 104 CTGAGGCGCATGCCCTCCCGAGGACCGTCTGCAGCCCTCTGCTCCGCGATGCTCTGGC 163
Db : : : : :
72 GlnProProProProGlnSerThrSerProProProValAlaThr----- 87
QY 164 TGGACTTGGCCATGGCAGGCTTCTGAGCCCTGAACACACAGAGAGTCCAGGTGA 223
Db : : : : :
88 ---ThrProAlaLeuProProLys-ProLeuPro---ProProLeuSerProPro 105
QY 224 GACCTCCCAAGCCCAACATGTTGTCAGCCCT----- 260
Db : : : : :
105 nThrThrProProProProAlaIleThrProProProProAlaIleThrProPr 125
QY 261 -----GCCACTTAGCAACAGC 277
Db : : : : :
125 oLeuSerProProProAlaIleThrProProProProAlaIleThrProProAl 145
QY 278 TCTGTGACCTGGAGCAGCGCCATCTCTGGGCTTCAGTCTTCTCCAGAGCACAAGG 337
Db : : : : :
145 aLeuProProLysProLeuProProLeu----- 155
QY 338 ACTCTGGGTCTGACCTCAGTCTTCTGGAAGGACATGGGGCTTAGAGTCTTAAACAGAC 397
Db : : : : :
156 -----SerProProGlnTh 160
QY 398 TGTTCCTCCCTTCCAGCAGAGAAGGAGTCGAAGAGCCACAGCCAGCTGCAGCCC-- 455
Db : : : : :
160 rThrProPro-----ProProProAlaIleThrProProLe 172
QY 456 -----CGAGCTCTAGCAGG 469
Db : : : : :
456 -----CGAGCTCTAGCAGG 469

Best Local Similarity: 30.00% Mismatches: 52
Query Match: 9.37% Indels: 37
DB: 2 Gaps: 10

US-10-659-782A-11 (1-579) x B56201 (1-506)

```
QY 62 ACCACTCCGCCAGGACTGCAG--GCCACCTGTCTGCAACCCAGCTGAGGCGCATGCCCT 118
DB 343 SerSerProAlaGlyLeuAlaLeuGlyAsnValThrAlaTrpGln---GlnProGlnPro 361
QY 119 CCCAGGACCGCTCTGCAGCCTCTCTGCTCCGCGCATGCTCTGGCTGGAGCTTGGCCATGG 178
DB 362 ProGlnProGln-ProProGlnProProGln----- 372
QY 179 CAGGCTCCAGCTCTCTGAGCCCTGAAACACAGAGAGTCCAGGTGAGACCTCCCCCAAAAG 238
DB 373 -----SerGlnProGlnProProGlnProGln---ProGlnGln-ProProGlnGlnG 389
QY 239 CCCACATGTTCTCCAGCCCTGCCACTTAGCAACACAGCTCTGTGACCTGGAGCAGCAGC 298
DB 389 InProHisLeuValPro---ValSerLeuSerAsn---LeuLeu---ProGlySerProLe 406
QY 299 GCATCTCTGGGCTTCAGTCTT---CTCCAGAGCACAAGAGGACTCTGGGTCTGACCTCA 355
DB 406 uProHisValGlyAlaAlaLeuThrValThrHis-----ProHi 420
QY 356 CTGTTTCTGAGGACATCGGGGCTTAGAGTCTTAACAGACTGTTTCCCTCCAGCA 415
DB 420 sileSerIleLys-----SerGluProValSerProSerArgl 433
QY 416 GAGAAAGAGTTCGAGAGCCACCCAGCC-----AAGCTGCAGCCCGGAGGTGAA 466
DB 433 uArgSerProAlaProProProProAlaValPheProAlaAlaArgProGluProGlyGl 453
QY 467 AGGCTGGCTCCCGCCGAGAGATGGAGGTCAAGCAGAGGGGCGAGAGGTAA 518
DB 453 uGlyLeuSerSerProAlaGlyGlySerTyroGluThrGlyAspArgAsp 470
```

RESULT 15

A56201
transcription factor MEF2D isoform 1a - mouse
C:Species: Mus sp. (mouse)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 15-Mar-2004
C:Accession: A56201
M:Martin, J.F.; Milano, J.M.; Hustad, C.M.; Copeland, N.G.; Jenkins, N.A.; Olson, E.N.
Mol. Cell. Biol. 14, 1647-1656, 1994
A:Title: A Mef2 gene that generates a muscle-specific isoform via alternative mRNA splicing
A:Reference number: A56201; MUID:94158837; PMID:8114702
A:Accession: A56201
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-514 <MAR>
A:Cross-references: GB:S68893; NID:G545519; PIDN:AAB29973.1; PID:G545520
C:Genetics:
A:Gene: Mef2d
C:Keywords: alternative splicing; transcription factor
F:2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Alignment Scores:

Pred. No.:	1-62	Length:	514
Score:	96.50	Matches:	48
Percent Similarity:	44.38%	Conservative:	23
Best Local Similarity:	30.00%	Mismatches:	52
Query Match:	9.37%	Indels:	37
DB:	2	Gaps:	10

US-10-659-782A-11 (1-579) x A56201 (1-514)

```
QY 62 ACCACTCCGCCAGGACTGCAG---GCCACCTGTCTGCAACCCAGCTGAGGCGCATGCCCT 118
DB 351 SerSerProAlaGlyLeuAlaLeuGlyAsnValThrAlaTrpGln---GlnProGlnPro 369
QY 119 CCCAGGACCGCTCTGCAGCCTCTCTGCTCCGCGCATGCTCTGGCTGGAGCTTGGCCATGG 178
```

```
DB 370 ProGlnProGln-ProProGlnProProGln----- 380
QY 179 CAGGCTCCAGCTTCTGAGCCCTGAACACACAGAGAGTCCAGGTGAGACCTCCCCCAAAAG 238
DB 381 -----SerGlnProGlnProProGlnProGln---ProGlnGln-ProProGlnGlnG 397
QY 239 CCCACATGTTGTTCCAGCCCTGCCACTTAGCAACACAGCTCTGTGACCTGGAGCAGCAGC 298
DB 397 InProHisLeuValPro---ValSerLeuSerAsn---LeuLeu---ProGlySerProLe 414
QY 299 GCATCTCTGGGCTTCAGTCTT---CTCCAGAGCACAAGAGGACTCTGGGTCTGACCTCA 355
DB 414 uProHisValGlyAlaAlaLeuThrValThrHis-----ProHi 428
QY 356 CTGTTTCTGAGGACATCGGGGCTTAGAGTCTTAACAGACTGTTTCCCTCCAGCA 415
DB 428 sileSerIleLys-----SerGluProValSerProSerArgl 441
QY 416 GAGAAAGAGTTCGAGAGCCACCCAGCC-----AAGCTGCAGCCCGGAGGTGAA 518
DB 441 uArgSerProAlaProProProProAlaValPheProAlaAlaArgProGluProGlyGl 461
QY 467 AGGCTGGCTCCCGCCGAGAGATGGAGGTCAAGCAGAGGGGCGAGAGGTAA 518
DB 461 uGlyLeuSerSerProAlaGlyGlySerTyroGluThrGlyAspArgAsp 478
```

Search completed: February 1, 2005, 13:29:29
Job time : 39 secs

THIS PAGE LEFT BLANK